

1. SCIENTIFIC EXPLANATION AND UNDERSTANDING

TIM THORNTON

International School for Communities, Rights and Inclusion
University of Central Lancashire

ABSTRACT

Jaspers's binary distinction between understanding and explanation has given way first to a proliferation of explanatory levels and now, in John Campbell's recent work, to a conception of explanation with no distinct levels of explanation and no inbuilt rationality requirement. I argue that there is still a role for understanding in psychiatry and that is to demystify the assumption that the states it concerns are mental. This role can be fulfilled by placing rationality at the heart of understanding without a commitment to the attempt to use rationality to shed light on interpretation and mindedness as though from outside those notions. Delusions still present a significant challenge to philosophical attempts at understanding, but this merely reflects the genuine clinical difficulties such states present.

Keywords: John Campbell, interventionism, normativity, understanding, Ludwig Wittgenstein

1. Levels of explanation in the history of psychiatry

A century ago, Karl Jaspers, influenced by the long running *Methodenstreit*, argued that explanation in psychiatry should be complemented by understanding. He was not hostile to explanation but, at a time when recent advances in neurology had led psychiatry to be dominated by what Jaspers called 'brain mythologists', he wished to balance that natural science emphasis with a human science element.

Jaspers's distinction is epistemic rather than ontological. Understanding and explanation do not have two distinct subject matters but differ in the method or the kind of intelligibility that they deploy. As applied in psychiatry, they share the same subject matter: 'real events' or 'thought processes', which can in principle be charted in either way.

The suggestive assumption that the psychic is the area of meaningful understanding and the physical that of causal explanation is wrong. There is no real event, be it physical or of psychic nature, which is not

accessible to causal explanation...

The effect a psychic state may have could in principle lend itself to a causal explanation, while the psychic state itself of course must be phenomenologically (statically) understood. It is not absurd to think that it might one day be possible to have some rules which could causally explain the sequence of meaningfully connected thought processes without paying heed to the meaningful connections between them ... It is therefore in principle not at all absurd to try to understand as well as to explain one and the same real psychic event. These two established connections, however, are of entirely different kinds of validity. (Jaspers 1974, 86)

Whilst explicitly contrasting explanation and understanding has fallen out of fashion, a distinction akin to Jaspers's remains influential in some philosophical circles. Wilfrid Sellars distinguishes what he calls the 'manifest image of man in the world' from the scientific image (Sellars 1963, 6). Donald Davidson argues that the mental is governed by the 'constitutive ideal of rationality' which has 'no echo in physical theory' (Davidson 1980; 223, 231). Building on Sellars's work, John McDowell contrasts the logical space of reasons with the realm of law or of natural science (McDowell 1994).

Whilst these distinctions are less clearly epistemological than Jaspers's, they are essentially tied to an epistemological distinction. The 'manifest image', the 'constitutive ideal of rationality' and the 'space of reasons' are all characterised in *normative* terms. And thus all are subject to a particular kind of intelligibility which can, for convenience be mapped onto the understanding side of the distinction between understanding and explanation. Understanding depends on a comparison with an ideal whilst explanation does not. According to a broadly Hempelian model, explanation depends on subsumption under a generalisation. According to a more recent interventionist view (see below), it depends on stable possibilities of intervention. But, however explanation is construed, it has a different kind of logic from understanding: paradigmatically how things usually happen, rather than how they ought to happen.

In psychiatry, there is an important further way in which normative understanding may play a central role. If the very idea of an illness, disease or disorder is a normative notion and if this cannot be reduced to more basic non-normative notions then assessing a subject's condition will also involve something like comparison to an ideal. But given that there are ongoing attempts to reduce the concept of disorder, at least, to failure of biological function itself construed as discharging any unexplained normative notions, this is a more contentious case (Wakefield 1999).

Whilst the firm distinction between normative and non-normative remains influential in some quarters, it is less so in much contemporary philosophy of psychiatry. There are two key indications of this. The first is that Jaspers's duality of understanding

and explanation has been replaced by a multiplicity of explanatory levels or kinds of approach. In his introduction to the philosophy of psychology, José Bermúdez takes the key challenge for this area of philosophy to be the relation between a number of different levels of explanation:

commonsense psychological explanation ... explanations of cognition and mental operations given scientific psychology, cognitive science, cognitive neuroscience and the other levels in the explanatory hierarchy (Bermúdez 2005, 35)

More specific to *psychiatric* explanation, Dominic Murphy, for example, has recently attempted to show how Marr's threefold distinction between levels can be applied to psychiatric explanation. Marr's threefold distinction is between:

Computational theory: What is the goal of the computation, why is it appropriate, and what is the logic of the strategy by which it can be carried out?
Representation and algorithm: How can this computational theory be implemented? In particular, what is the representation for the input and output, and what is the algorithm for the transformation?
Hardware implementation: How can the representation and algorithm be realized physically? (Marr 1982, 25)

Murphy points out that Marr thinks of the levels in epistemic terms: as different ways of understanding the same system. One can determine its goals, the algorithm by which it determines those goals or the physical set up which implements that algorithm. But one might think of them as describing distinct forms of organisation in nature or distinct causal structures pitched at different ontological levels: 'higher levels are made up of lower level things, and at each level things interact with each other rather than with things at lower levels' (Murphy 2008, 10).

In psychiatry, however, neither of these pictures is quite right, according to Murphy, because there are different systems operating at different levels, unlike the epistemic view, but the different levels interact, unlike the latter view. Thus whilst Marr's description of levels suggest that they are partly independent (the computational level constrains but does not determine the causal mechanisms that implements it), Murphy suggests that in psychiatry causes described at one level will have effects at another so that useful generalisations will cross levels.

The replacement of an opposition between Jaspers's two levels with a proliferation is merely the first of three steps that have reduced the role of understanding in psychiatry. The second is the pragmatic attitude to switching between levels described by Murphy. Indeed, the dominant view is that 'most explanations in psychiatry will end up being cross levels rather than solely within levels' (Kendler and Parnas 2008, 100). But the third step is, I suggest, provided by a series of articles by John Campbell in which he both argues against the very idea that there are levels of explanation but also that the model of causal explanation that remains has no need of intelligibility.

2. Campbell on causation and explanation in psychiatry

Campbell argues that the multi-level model of explanation in psychiatry results from a pre-Humean assumption about the intelligibility of relations tracked in causal explanation. He gives, as an example, a discussion of thought insertion by Christopher Frith. Frith claims that whether or not inappropriate firings of dopamine neurons are found in subjects who experience thought insertion, that fact cannot be used to explain their experiences as it would shed no light on why that kind of symptom, rather than another, was produced by such firing. To shed light, Frith assumes, we need an account pitched at a particular level: in Frith's case that of a sub-personal but still cognitive model of mechanisms supposedly responsible for thought insertion.

Campbell suggests that the assumption that there is a right level of explanation which clarifies things in the way Frith desires is the result of a pre-Humean view of causal explanation. Although it is easily forgotten, Hume taught us that there need be no intelligible connection between cause and effect. That is implicit in his rejection of any logical connection to analyse the necessitating relation between them. Causal connections are merely brute facts to be discovered by experience.

Resisting the idea that the right kind of cause and effect have to be intelligibly, rather than merely brutally, related also undercuts the motivation for the levels of explanation picture.

We naturally seek a certain kind of intelligibility in nature; we naturally try to find explanations that will show the world to conform to reason, to behave as it ought. Hume's point is that there are no such intelligible connections to be found. This point has generally been accepted by philosophers thinking about causation... Hume's comments nonetheless do leave us in an uncomfortable position, because we do tend to look for explanations that make the phenomena intelligible to reason. We are prone to relapse, to think that after all we must be able to find intelligibility in the world. This tendency survives, I suspect, in the idea of 'levels of explanation'. The idea is that within certain levels of explanation, we will find a particular kind of intelligibility... [T]he lesson from Hume is that there is no more to causation than arbitrary connections between independent variables of cause and effect. We have to resist the demand for intelligibility. (Campbell 2008, 201)

This is not just a restatement of the recent pragmatic attitude that, in psychiatry, explanations may cross levels. Rather, the very idea of levels of explanation, where these are understood as characteristic forms of intelligibility, is itself undercut.

This leaves, however, the issue of shedding some light on the nature of causal connections (if not a priori light on particular causal connections). In (non-mental) cases of causation, the notion of *mechanism* plays a central role in empirical research. Searching out the way in which causal influence is transmitted has been an important

part of scientific practice. ‘It would seem a kind of madness if someone were to acknowledge that there is a causal link, but propose that there may be no mechanism linking the two’ (Campbell 2009, 138). But if science has usefully explored the mechanisms that mediate causal influence, there must be some paradigmatic mechanisms that stand in need of no further explanation and the transmission of motion by impulse, in Hume’s billiard ball example, is one such prototype.

Nevertheless, the idea that there *must* be such a mechanism is a kind of synthetic a priori claim which, Campbell suggests, should be rejected in line with Hume’s argument. He adopts the interventionist model defended most extensively by James Woodward in *Making Things Happen* according to which for X to be a cause of Y is for intervening on X to be away of intervening on Y (Woodward 2003). The rejection of the necessity of a mechanism and the adoption of an interventionist approach opens up the possibility of a causal connection – in accord with interventionism – but where there is no mechanism. In the case of psychiatry, however, the key issue is causation in the absence of a *mental* mechanism, whatever that is taken to be.

Just as we find it natural to expect there to be a mechanism underpinning material causal connections – even if this assumption lacks any genuine a priori justification – so Campbell also suggests that in the case of mental causation we expect there to be a rational connection between propositional attitudes. The rational link between two propositional attitudes is our paradigm of a mental causal mechanism. So if one hears someone explain that they believe that Tranmere Rovers won their most recent football match because they heard it on the BBC, which they take to be trustworthy, no further inquiry is needed as to why the beliefs about what they hear and trust cause the belief about the result. Again, however, whilst the idea that mental causation is underpinned by rational connections is natural and compelling, it lacks a priori justification.

[T]here is an analogy between:

- 1 the idea that propositional attitude ascriptions depend on the ascription of rationality to the subject, and
- 2 the idea that all causal interactions between pieces of matter must be comprehensible in mechanistic terms.

Both ideas express an insight – that we find it extremely puzzling when we encounter causal relations among propositional attitudes that are not broadly rational, just as we find it extremely puzzling when we encounter causal interactions between physical objects that are not mechanistic, and that involve spooky ‘action-at-a-distance’. Both ideas express a natural impulse of philosophers – to elevate this kind of point into a kind of synthetic a priori demand that reason makes on the world. This impulse has to be resisted. (Campbell 2009, 142)

So in both cases there is a genuine insight. As a matter of custom and habit, we find an absence of material mechanisms and an absence of rational connections between

mental states puzzling. But in both cases it is a characteristic philosophical error to promote this natural expectation into a justified a priori claim that the world must respect. Mere custom and habit cannot rationally sustain any such demand on how the world must be.

The rejection of the necessity for rational connections between causally related mental states looks to ease a central problem for the philosophy of psychiatry: explaining delusions. There need be nothing genuinely mysterious about a causal connection which lacks a rational connection (the expected mental mechanism).

Suppose you believe:

- 1 that this man is stroking his chin, and
- 2 that this man believes you need to shave.

What is it for the first belief to be a cause of the second? On the interventionist analysis, it is for the intervention on the first belief to be a way of changing whether you have the second belief. So if some external force changed your belief that this man is stroking his chin, you would no longer believe that he believes you need to shave. There is no appeal to rationality here, no appeals to mechanism. (Campbell 2009, 143)

The causal connection between one state and another is underpinned in interventionist terms based on the idea that if intervening on the first belief is a stable way of bringing about a change in the second then this is sufficient for there to be a causal connection between them.

Spelling this idea out involves a little more complexity, however. Given a scanner capable of yielding a complete microphysical description of the human body and a longitudinal study of schizophrenia in a population, Campbell suggests that it might be possible to form a disjunctive characterisation of the set of microphysical states that are nomically sufficient for schizophrenia. But that function from physical states to illness would lack any concise expression and would not be couched in terms of variables which could be affected by local processes.

For propositional attitudes to count as causes of delusions, Campbell suggests two conditions have to be met. There should be 'systematic relations between cause variables and the subsequent delusion' and there should be a correlation between a change of the cause and a change of the effect (Campbell 2009, 146). More generally for the causal explanation of mental states, the causal variables, which he calls 'control variables', should have large, specific and systematic correlations with their effects akin to the way the controls of a car systematically control its behaviour. These conditions do not require a rational connection, however. To repeat Campbell's phrase, there need be 'no appeal to rationality here'.

The classical philosophical approach has been to regard propositional attitudes as part of a 'conceptual scheme' that we bring to bear in describing

the ordinary world. This conceptual scheme is taken to have strong a priori constraints on its applicability. In particular, as we have seen, rationality is taken to be a norm with which the scheme has to comply... The appeal I have just been making to the notion of a control variable is intended to replace this invocation of rationality... [I]t is the fact that we have control variables, not the fact that we have rationality, which means that we are 'at the right level' to talk of beliefs and desires. (Campbell 2009, 147)

The phrase 'at the right level' occurs in inverted commas to flag the fact that the notion of the right explanatory level has been undercut. Without a pre-Humean insistence on the intelligibility of causal relations there is no more to the notion of being at the right level than that there is a causal relation tracked through the idea of control variables.

With the idea of control variables replacing an a priori requirement for rationality in mental causation, psychiatric explanation of delusions is in principle in the same predicament as the explanation of any other belief. Causal explanation has been achieved once one has an understanding of the variables necessary for changing the delusional belief entertained. The apparently principled problem of attempting to fit primary delusions into some sort of rational framework is replaced by a practical problem of charting the variables that affect them. But is that aim enough for psychiatry?

3. A continuing role for understanding?

As I have summarised above, Campbell's approach to causal explanation within psychiatry has two aspects: the rejection of the synthetic a priori framework, which would impose a rational structure on the causal explanation of mental states including delusions, and an austere interventionist approach to causation itself. The argument against the synthetic a priori framework is Humean. There is no reason to think that cause and effect stand in an intelligible, rather than a merely brute, relation. Thus, although we tend to assume that the paradigm of mental causation is a rational connection between mental states – just as we tend to assume that the paradigm of material causation is the communication of motion by impulse – there is no reason that this must be so. It is a temptation we, as philosophers, should resist.

Whilst that might be the case as far as causation is concerned, it does not explain why the states invoked as control variables should count as *mental* states. This is not to say that it conflicts with or precludes that idea. Rather, the account of control variables – which according to Campbell replaces an appeal to rationality – is insufficient to make that characterisation clear. Slimming the resources of explanation down merely to a causal interventionist account leaves the status of the mental mysterious. Whilst there are good reasons to hold that *intelligibility* is no part of the causal relation per se, there are good reasons, which flow from twentieth century criticisms of Cartesianism, to hold that it is part of what it is to be mental. Given that it is not a part of the

interventionist explanation of mental states itself, intelligibility is a necessary adjunct to that.

But it is one thing to reject the idea that minds might be essentially private and stress instead their necessary intelligibility, it is another to say something about what that intelligibility comprises. To say something about that is to sketch out the shape that a form of psychiatric *understanding* can take in addition to the austere conception of interventionist causal *explanation* that Campbell outlines (to adopt that familiar distinction). Like Jaspers's aim, it is to sketch a continuing role for understanding in addition to, and despite recent developments in, the causal explanation of mental disorder. (But this is not to say that the schematic discussion of understanding that follows mirrors Jaspers's own account of understanding, of phenomenology and empathy.)

The most promising recent approach to shedding some light on the nature of such intelligibility has been to explore the connections between interpretation, mentality and rationality. On this approach, the interpretation of other minds presupposes that they fit into a broadly rational pattern. The most influential version of this approach – Donald Davidson's – bases the connection between mindedness and rationality on an explicit argument (Davidson 1980). The justification of any ascription of mental states to another person is limited in principle to those available to a radical interpreter: i.e. someone interpreting from scratch without the help of resources such as dictionaries or bilingual speakers. Whilst this does not describe the less austere predicament of everyday interpreting, it is a rational reconstruction of all that could ultimately be offered as justification for any interpretation.

Davidson argues, however, that the holism of beliefs and meanings presents a fundamental problem for interpretation. Only if one already knows the belief expressed in an utterance can one translate the utterance. And only if one can already translate the utterance can one determine the belief expressed. It is to break into this circle that Davidson invokes rationality. One must assume that beliefs are broadly true and rational: an assumption called 'the Principle of Charity'. The assumption of rationality is thus an essential feature of understanding other minds. And, given that radical interpretation outlines in principle all the facts available to determine meaning and mental content, this implies that they are constitutively rational. Hence Davidson outlines an argument for the claim that understanding has a distinctive rational shape.

There are two main kinds of objection to Davidson's argument that are relevant here. First, there are arguments that Charity is unnecessary for interpretation. Weaker principles such as the Principle of Humanity might do the job better. Second, it has been argued that psychiatric symptoms such as delusions are counter examples to the claim that the mental is constitutively rational (Bortolotti 2005). Before considering these objections, however, it will be helpful to clarify the role that rationality need play in shedding light on mindedness.

One need not assume that rationality is supposed to shed light on the nature of mind and meaning from outside. If the connection between rationality and mentality were to provide an independent insight into the latter it would be necessary to have a characterisation independent of the intelligibility of minds in interpretation. (One indication of this aim in Davidson's own thinking is his suggestion that truth can be used to shed light on meaning, inverting Tarski's order of explanation (eg. Davidson 1980, xiv). But an indication that he may not intend such an analysis is his later suggestion that the structure and content of truth itself is the pattern exemplified in radical interpretation (Davidson 1990).)

But the connection between mindedness and rationality need not be made as an attempt to shed light on the former from an external position. Perhaps the best insight we can have into the nature of rationality is manifested in the pattern of intelligibility tracked in radical interpretation.

This approach is suggested by two further considerations. The first is that the way that the rationality of the mental is exemplified by the normativity of mental content and linguistic meaning itself. As Wittgenstein emphasises, mental states stand in prescriptive relations to the states and events that satisfy them whilst an understanding of meaning is an understanding of the correct application of words.

A wish seems already to know what will or would satisfy it; a proposition, a thought, what makes it true – even when that thing is not there at all! Whence this *determining* of what is not yet there? This despotic demand? (“The hardness of the logical must.”) (Wittgenstein 1953, §437)

Such normative connections are intrinsic features of the mental, constituting states as the states they are, rather than an external way of explaining or characterising the mental from without. One of the key themes of the *Philosophical Investigations* is that no attempt to explain the normativity of the mental in other terms can succeed. Normativity is also a feature of linguistic meaning, of, for example, Davidson's recommended codification of the output of radical interpretation: instances of the T-schema.

[W]hat makes it *correct* among speakers of English to make a claim with, say, the words ‘Snow is white’... is that snow is indeed white. I stress ‘correct’: truth in the sense of disquotability... is unproblematically normative for the practice of using the sentence mentioned on the left-hand side of T-sentences... (McDowell 2009, 214)

Such instances display the normativity of linguistic meaning and thus partially exemplify the rationality Davidson invokes in radical interpretation. But, again, there is no need to assume that their normativity can be cashed out in non- or pre-linguistic terms.

On the Tarskian-Davidsonian conception the ‘oughts’ in question – the ‘oughts’ that are built into the idea of, say, denotation – are not separable from the idea of correctness in assertion... I think once we see that the intuition that meaning and aboutness are ‘ought’-laden does not require the relevant ‘oughts’ to be pre-semantic... we can see that there is no ground for the idea that linguistic behaviour must be governed by... proprieties that can be formulated in non-semantic terms... (McDowell 2009, 215-6)

The second consideration to support the claim that the connection between mindedness and rationality need not be made as an attempt to shed light on the former from an external position stems from consideration of interpretation of the speech and action. According to the approach I am considering, the paradigmatic pattern of such interpretation is manifestly different from the explanation of other natural phenomena by, for example, subsumption under laws. (To be merely *broadly* Davidsonian it is not necessary to insist that all action explanation implicitly relies on a practical syllogism pairing beliefs and desires.)

Both involve placing things in a pattern. But in one case the pattern is constituted by regularities according to which phenomena of the relevant kind unfold; in the other it is the pattern of a life led by an agent who can shape her action and thought in the light of an ideal of rationality. (McDowell 2009, 211)

The ideal of rationality appealed to here is one appropriate to the ‘pattern of a life led by an agent’. There is an important contrast with the idea of subsumption under a regularity, of what generally happens. Instead, the interpretation turns on the normative idea of what is *right* for agent. But again, there is no need for this point to imply that the rationality implicit in interpretation of speech and action can be unpacked in terms independent of intelligible connections between mental states, utterances and actions.

So it is possible to take a broadly Davidsonian approach, which places a rational structuring at the centre of understanding speech and action, without committing oneself to an independent characterisation of rationality. If so the dispute about the rival claims of the principles of Charity and Humanity is not relevant here. Without the assumption that the description of the presuppositions of radical interpretation are supposed to shed light on mindedness as from outside, little hangs on the distinction between them.

What of the objection that delusions are counter examples to the claim that the mental is constitutively rational? In rejecting the role of synthetic a priori claims about intelligibility, Campbell mentions and dismisses Davidson’s approach on the grounds that delusions look to be mental states but do not readily fit into a rational structure. As Campbell says, it ‘stretch[es] credulity that the figure of the schizophrenic could be used to explain what Davidson meant by “rational”’ (Campbell 2009, 141; cf Bortolotti 2005).

This point does indeed present a difficulty for any approach which assumes that rationality is key to intelligibility and that intelligibility is of the essence of the mental. Nevertheless, just as hard cases make for bad laws, the problem of interpreting delusions is not reason enough to give up an interpretation-based approach to the mind as a whole, not least because making sense of the content of delusions can be a clinical as well as a philosophical problem. (Thus any philosophical account which failed to save that phenomenon – the clinical difficulty – would be undermined rather than reinforced by that fact.) So the fact that subjects of schizophrenic delusions, or perhaps the delusions themselves, are far from paradigmatic instances of rationality – and thus would not be used to teach the meaning of ‘rationality’ – need not undermine the central thesis. In the next section I will try to say something more about this.

4. Understanding applied to delusions

To outline the general strategy I commend, I will briefly consider a recent criticism of the idea that understanding must have a rational shape made by Lisa Bortolotti (Bortolotti 2005). She outlines two broadly Davidsonian strategies for accommodating delusions within a rational framework. One is to concede that delusions are irrational in a different (more dramatic) way than other beliefs and thus to deny that they are beliefs at all. The second is to argue that delusions are beliefs and are irrational in the same way in which other beliefs can be: amenable to revision and correction. She then outlines objections to both strategies. The problem with the first move is that delusions behave like beliefs (for example, they guide action). The problem with the second move is that delusions are not as easy to revise and correct as she suggests broadly rationalist approaches require. So, the puzzle with monothematic delusions is that they behave like beliefs but are not largely rational on what she takes to be the most charitable approach to that idea. It is the second argument I will consider here.

Bortolotti targets what she calls the ‘Background Argument’ which assumes that occasional and local breakdowns of rationality can only be accommodated within an ascription of mindedness if there is a background of rationality. [O]nly within a belief system that is largely rational can intentional description occasionally be granted to representational states that fail to meet the standards of rationality’ (Bortolotti 2005, 189). She goes on to characterise the rational constraints on interpretation that the Background Argument imposes more specifically:

There are two main requirements that a creature A must satisfy in order to count as a believer: (1) A cannot explicitly violate a fundamental norm of rationality and (2) A cannot fail to recover from a violation of a norm of rationality once this violation is made explicit to A. This capacity to restore rationality, together with the general conformity of the creature’s behaviour to the standards of rationality, is regarded as sufficient for some version of RC [the rationality constraint] to be in place. (Bortolotti 2005, 190)

Her prima facie case is then that monothematic delusions such as Capgras do not meet this modest remaining condition. Two arguments for this conclusion are contained in the following lengthy passage.

The only reading of [the Background Argument] BA that would make it true that monothematic and circumscribed delusions can be intentionally characterised is the reading according to which a belief system is largely rational if most of the beliefs it contains are rational. This reading lacks the resources to discriminate between the behaviour of believers who can recover from failures of rationality and those who cannot, a distinction that plays a very important role in Davidson's characterisation of BA. The simplistic reading cannot therefore be the best available reading of BA. It is not consistent for the Davidsonian to illustrate the role of BA by appealing to the phenomenon of recovery and to endorse the view that deluded subjects are largely rational. For Davidson, it is a condition on believers that, after 'Socratic tutoring', they recover from deviations from rationality. Deluded subjects typically don't.

There is a further reason why I believe it is not in the spirit of BA to consider the behaviour of people affected by relatively circumscribed and monothematic delusions as largely rational. The idea behind BA is that the non-rational belief can be made sense of only by appealing to the other true and sensible beliefs the subject has. That a false belief can be explained by appealing to the other true and justified beliefs that belong to the same system, is often true... The same strategy cannot be easily applied to delusional states. Which true belief of Bob will help us rationalise his conviction that the person who looks identical to his mother is actually an alien? (Bortolotti 2005, 203-4)

The first argument turns on the fact that sufferers of monothematic delusions differ from people who make specific logical errors. The latter typically can be taught to recognise and correct their beliefs whereas as the former typically hold onto their delusional beliefs come what may. So if the claim that subjects must be largely rational is defended for the case of logical errors through the notion of recognition of error and self-correction, that cannot help in the case of delusions. Second, the assumption of rationality is typically reconciled to a subject holding some false beliefs through an explanation using true beliefs but that cannot help in the case of monothematic delusions.

The quick response to both these arguments is to concede that the methods that reconcile an assumption of necessary rationality with logical and other non-psychopathological errors will not apply in the case of some delusions, at least. That fact is taken by Bortolotti to undermine the rationalist approach to mentality. But it need not. Since monothematic delusions are pre-philosophically genuinely puzzling whilst logical errors are not pre-philosophically puzzling it would count *against* an

account of understanding if it were to produce the same results in both cases. (So this response claims that the sophisticated reading of the Background Argument, according to which rationality requires agents to amend and correct errors to count as agents at all, is too strict. The slow response requires a painstaking approach to each individual case to see what rational relations hold, and what do not, without overly prejudging the shape that will take.)

Take the case of Bob who supposedly believes that the person who lives in his house and looks identical to his mother is not his mother but a cleverly disguised Martian. How seriously should we take this belief ascription (helpfully provided by Bortolotti)? The problem is this. Bob does not have the supporting beliefs that we would have were we to believe that someone – almost incredibly under any circumstances – had been replaced by a cleverly disguised Martian. Given that, then the ascription of beliefs to him comes under (pre-philosophical) strain. Does he really believe that the imposter is a Martian or just an alien called for convenience in his idiolect a ‘Martian’? What supporting discriminating beliefs does he have about the solar system? To the extent to which these questions can be answered then the assumption of rationality looks to be the best tool to chart his conception. But that is not to say that these questions can be answered. That is the puzzle of delusions and the reason Jaspers called them ‘un-understandable’.

To what extent, then, can one bring understanding successfully to bear on delusions? I think that this is a genuine challenge for both psychiatry and philosophy of psychiatry alike and will not attempt a general answer to that question here. But even within an approach that takes rationality to be central to mindedness there are more ways to attempt to gain some understanding than at first appear. The clue to one such approach is to be found in some of Wittgenstein’s discussion of meaning and understanding in cases which do not fit within the domain of everyday shared practices or uses.

First, it is necessary to avoid the pitfalls of both platonism and constructionism about meaning. Wittgenstein opposes a notion of meaning that severs all connection to our meaning-laden reactions and instead assumes platonic rails leading to infinity. Thus, by extension, he opposes a picture in which one can begin to gesture towards other meaning-driven ways of going on, of which, however, we can make nothing – cannot understand or follow – but can still think of, from a detached philosophical perspective at least, as norm- or meaning-driven (cf Lear 1982, 1984, 1986).

At the same time we should not react to this rejection by adopting a view of meaning in which we simply make things up as we go along. We should not think of meaning as a matter of piecemeal construction since that undermines its necessary normativity (contra Kripke 1982, Wright 1986). The familiar connection of meaning and practice need not be taken to reduce meaning to a notion of practice that could be understood in non-normative, non-meaning-related terms.

Notwithstanding the rejection of platonism – and thus the rejection of the idea, at least,

of other ways of going on – this does not imply that where rational interpretation fails we must have fully left the domain of meaning and our meaning-related reactions. It is possible to maintain the rejection of platonism but still have a more pluralistic response to its absence.

A useful approach is to think about an ‘intransitive’ notion of meaning. We can think of this on the basis of what Wittgenstein says in passages of the *Brown Book* in which he criticises a constitutive theory of meaning and distinguishes transitive and intransitive uses of ‘particular’ in order to criticise the idea that meaning something consists in a ‘particular’ experience. (Nothing hangs on the potentially misleading use of the word ‘delusion’ here.)

Now the use of the word “particular” is apt to produce a kind of delusion and roughly speaking this delusion is produced by the double usage of this word. On the one hand, we may say, it is used preliminary to a specification, description, comparison; on the other hand, as what one might describe as an emphasis. The first usage I shall call the transitive one, the second the intransitive one. Thus, on the one hand I say “This face gives me a particular impression which I can’t describe”. The latter sentence may mean something like: “This face gives me a strong impression”. (Wittgenstein 1958, 158)

The intransitive notion of meaning looks to be in play in a passage in which Wittgenstein compares musical and semantic understanding such as:

Understanding a sentence is much more akin to understanding a theme in music than one may think. What I mean is that understanding a sentence lies nearer than one thinks to what is ordinarily called understanding a musical theme. Why is just *this* the pattern of variation in loudness and tempo? One would like to say “Because I know what it’s all about.” But what is it all about? I should not be able to say. In order to ‘explain’ I could only compare it with something else which has the same rhythm (I mean the same pattern). (One says “Don’t you see, this is as if a conclusion were being drawn” or “This is as it were a parenthesis”, etc. How does one justify such comparisons?--There are very different kinds of justification here.)

We speak of understanding a sentence in the sense in which it can be replaced by another which says the same; but also in the sense in which it cannot be replaced by any other. (Any more than one musical theme can be replaced by another.)

In the one case the thought in the sentence is something common to different sentences; in the other, something that is expressed only by these words in these positions. (Understanding a poem.)

Then has “understanding” two different meanings here?--I would rather say that these kinds of use of “understanding” make up its meaning, make up my concept of understanding. (Wittgenstein 1953 §§ 527, 531, 532)

These passages suggest that in addition to the everyday sense of meaning which applies to learning and using sentences to frame, for example, empirical judgements there is a further variant notion of meaning shorn of the connection to meaning something. (Cf. '[I]f, recognizing this, I resign myself to saying "It just expresses a musical thought", this would mean no more than saying "It expresses itself"' (Wittgenstein 1958, 166).) If so, how might this help with understanding psychopathology?

A recurring theme in the *Remarks on Frazer's Golden Bough* is Wittgenstein's criticism of Frazer's method of shedding light on the alien cultural practices by looking for underlying causes (such as quasi-scientific attempts to intervene in nature through rain dances) (Wittgenstein 1978). To a first approximation, Wittgenstein suggests instead that the meaning of cultural practices lies open to view (for those, at least, with eyes to see). One comparison he uses is to remind us of our habit of kissing a picture of one's beloved. This is not the means to any distinct end. Nor does it encode a particular thought. It is however, nevertheless, an aspect of symbolic use. A helpful suggestion made by Hutchinson, Read and Sharrock is that, they suggest, we are not supposed by Wittgenstein to find our own habit here unsurprising (Hutchinson, Read and Sharrock 2008, 47). Although this is what we do, still, on reflection, it is strange. Thus the comparison between the practices described by Frazer and our own habits does not so much provide a route to domesticate the former as show how we can begin to approach them with some degree of understanding without finding them ordinary.

The idea of intransitive meaning together with Wittgenstein's sketch of an alternative to Frazer's approach to alien rituals and practices suggests the beginnings of a way to deploy understanding in the face of delusions. On Frazer's own approach, the rituals described in one way mean something else. A dance that, *we* know, cannot have any effects on the weather, is nevertheless attempted for that reason. That is the kind of act it, intentionally, is. This style of interpretation is akin to Bortolotti's account of the Background Argument which, as she emphasises, is an implausible interpretation of monothematic delusions. (Thus Frazer's approach is an unpromising approach to monothematic delusions.) On Wittgenstein's approach, by contrast, we can come to an understanding of the rituals (albeit by a potentially painful process of self-transformation) in their own terms and without seeing them as the proper subjects for self-correction. Thus they lie outside the scope of the Background Argument. Nevertheless this is still a form of understanding even if not a form of translation into our practices. It takes place against a background of rational interpretation, of finding order in much of what the others say and do, even though it itself resists assimilation. Such understanding goes beyond the resources of merely causal explanation.

This is not to say that all or even any expressions of monothematic delusions are instances of intransitive meaning. My purpose is more modest but more general: to suggest that there can be space for a form of understanding, for interpretative work to be done, even when, as in the case of delusions, there are genuine pre-philosophical difficulties.

5. Conclusions

Jaspers argued that there was a particular need within psychiatry for understanding as well as explanation. That distinction has fallen from favour (both in thinking about psychiatry and within the philosophy of science more generally). Within the philosophy of psychiatry, a duality of explanation and understanding has been replaced by a variety of explanatory levels or approaches combined with a pragmatic attitude to their application. Explanation trading across levels is the expectation.

Very recently, John Campbell has suggested that the very idea of explanatory levels is a hangover from a pre-Humean view of causation which assumes that causal connections should have a degree of intelligibility (above brute causal connectedness). Campbell argues against any synthetic a priori assumptions of intelligibility of nature and for an austere interventionist model of causal explanation. In particular, this suggests a route for approaching delusions which can sidestep worries about their lack of rationality.

Despite the attractions of the interventionist view of causation, however, a psychiatry based solely on Campbell's suggestion lacks the resources for making the assumption that psychopathological experiences are still mental unmysterious. I have suggested that this is the role for understanding, in addition to explanation, in psychiatry. Furthermore, such understanding should assume the rationality of the phenomena it approaches without assuming that rationality can be characterised in terms independent of belief and meaning. The fact that delusions present a genuine challenge for a philosophically motivated approach to understanding does not undermine such an approach. Delusions present the genuine clinical puzzles that must challenge any model of understanding.

REFERENCES

- Bermúdez, J. L. 2005. *Philosophy of Psychology*. London: Routledge.
- Bortolotti, L. 2005. Delusions and the background of rationality. *Mind and Language* 20: 189-208.
- Campbell, J. 2008. Causation in psychiatry. In Kendler, K. S. and Parnas, J. (eds) *Philosophical Issues in Psychiatry*. Baltimore: Johns Hopkins University Press.
- Campbell, J. 2009. What does rationality have to do with psychological causation? Propositional attitudes as mechanisms and as control variables. In Bortolotti, L. and Broome, M. (eds) *Psychiatry as Cognitive Neuroscience*. Oxford: Oxford University Press.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davidson, D. 1980. The structure and content of truth. *Journal of Philosophy* 87: 279-328.
- Hutchinson, P., Read, R., and Sharrock, W. 2008. *There is No Such Thing as a Social Science*. Aldershot: Ashgate.
- Jaspers, K. (1913) 1974. Causal and "Meaningful" Connections between Life History and Psychosis, trans. by J. Hoenig, in S. R. Hirsch and M. Shepherd. In Hirsch, S.R., and M. Shepherd, *Themes and Variations in European Psychiatry*. Bristol: Wright: 80-93.

- Kendler, K. S. and Parnas, J. (eds) 2008. *Philosophical Issues in Psychiatry*. Baltimore: Johns Hopkins University Press
- Marr (1982).
- Kripke, S. 1982. *Wittgenstein on rules and private language*. Oxford: Blackwell.
- Lear, J. 1982. Leaving the world alone *Journal of philosophy* 79: 382-403.
- Lear, J. 1984. The disappearing 'we' *Aristotelian society supplement* 58: 219-58.
- Lear, J. 1986. Transcendental anthropology in Pettit, P. and McDowell, J. (eds) *Subject thought and context*. Oxford: Clarendon.
- Marr, D. 1982. *Vision*. San Francisco: W. H. Freeman.
- McDowell, J. 1994. *Mind and World*. Cambridge, Mass.: Harvard University Press.
- McDowell, J. 2009. *Having the World in View*. Cambridge, Mass.: Harvard University Press.
- Murphy, D. 2008. Levels of explanation in psychiatry. In Kendler, K. S. and Parnas, J. (eds) *Philosophical Issues in Psychiatry*. Baltimore: Johns Hopkins University Press.
- Sellars, W. 1963. Philosophy and the Scientific Image of Man. In *Science, Perception and Reality*. London: Routledge & Kegan Paul.
- Wakefield, J. C. 1999. Mental disorder as a black box essentialist concept. *Journal of Abnormal Psychology* 108: 465-72.
- Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.
- Wittgenstein, L. 1958. *The Blue and Brown Books*. Oxford: Blackwell.
- Wittgenstein, L. 1978. *Remarks on Frazer's Golden Bough*. Doncaster: Brynmill Press.
- Woodward, J. 2003. *Making Things Happen*. Oxford: Oxford University Press.
- Wright, C. 1986. Rule following meaning and constructivism. In Travis, C. (ed) *Meaning and Interpretation*. Oxford: Blackwell.

Received: April 15, 2010

Accepted: April 30, 2010

International School for Communities, Rights and Inclusion
University of Central Lancashire
Preston
PR1 2HE, UK
TThornton1@uclan.ac.uk